

Kodėl dirbtinis intelektas kelia egzistencinį pavojų žmonijai

Yuval Noah Harari

Istorikas, tarptautinio bestselerio „Sapiens“ autorius Yuval Noah Harari savo naujoje knygoje „Nexus“ nagrinėja, kaip per tūkstančius metų mūsų pasaulį formavo informaciniai tinklai. Ypatingą dėmesį autorius skiria dirbtiniam intelektui (DI). Jei atominių ginklų pavojus visiems suprantamas, suvokti DI pavojų sudėtingiau. Išskirtinėje knygos ištraukoje Y. N. Harari pasitelkia pramoninės revoliucijos istoriją paaiškinti DI iššūkius. Knyga visame pasaulyje, taip pat ir lietuviškai, pasirodys rugsėjo viduryje.

Ekspertai įspėja, kad nekontroliuojama DI pažanga gali baigtis didelio masto gyvybės ir biosferos praradimu, žmonijos marginalizacija ar net išnykimu. 2023 metų apklausoje, kurioje dalyvavo 2778 DI tyrėjai, daugiau nei trečdalis apklaustųjų nurodė bent 10 proc. tikimybę, kad pažengęs DI gali sukelti tokias blogas pasekmes kaip žmonijos išnykimas. 2023 metais beveik trisdešimties šalių, įskaitant Kinijos, Jungtinių Amerikos Valstijų ir Jungtinės Karalystės, vyriausybės pasirašė Bletchley deklaraciją dėl dirbtinio intelekto, kurioje pripažįstama, kad „dėl svarbiausių DI modelių galimybių kyla rimtės, net katastrofiškos tyčinės ar netyčinės žalos pavojus“.

Kai kurių žmonių nuomone, perspėjimai apie civilizacijos žlugimą skamba kaip perspausta jeremiada¹. Kiekvieną kartą, kai rasdavosi nauja galinga technologija, kildavo daug nerimo bangų dėl artėjančios apokalipsės, bet žmonija juk vis dar egzistuoja. Vykstant pramoninei revoliucijai, pražūtis scenarijus skelbė luditai, bet juk nė vienas iš jų nepasiteisino, o „tamsūs šėtoniški malūnai“, kuriais gąsdino Blake'as, galiausiai padėjo sukurti turtingiausias visuomenes istorijoje. Dabar dauguma žmonių gyvena nepalyginamai geresnėmis sąlygomis nei jų protėviai XVIII amžiuje. O dėl ateities, tokie DI entuziastai kaip Marcas Andreessenas ir Ray Kurzweilas žada, kad išmaniosios mašinos bus daug kartų naudingesnės už bet kokias ankstesniasias. Žmonės galės

¹ Jeremiada (pagal Biblijos pranašo Jeremijo vardą) – tai literatūrinis kūrinys, smerkianti kalba, pamokslas ir t. t., bibliinių pranašysčių tonu išreiškiantis nepritariamą ar įspėjantį apie nelaimę.

naudotis daug geresnėmis sveikatos priežiūros, švietimo ir kitomis paslaugomis, o DI net padės gelbėti ekosistemą nuo sunaikinimo.

Deja, atidžiau pažvelgus į istoriją paaiškėja, kad luditai ne taip jau ir klydo – iš tikrųjų turime labai rimtų priežasčių bijoti naujų galingų technologijų. Net kai teigiamos technologijų savybės galiausiai nusveria neigiamas, iki tokios laimingos pabaigos reikia įveikti daugybę išbandymų ir kančių. Dėl kokios nors naujos technologijos įvyko daug istorinių nelaimių, bet ne dėl jos inherentiško blogumo, o dėl probleminio naudojimo – žmonėms prireikia laiko išmokti išmintingai ja naudotis.

Pramoninė revoliucija yra puikus pavyzdys. XIX amžiaus pradžioje visame pasaulyje pradėjus plisti pramoninėms technologijoms, buvo suardytos tradicinės ekonominės, socialinės ir politinės struktūros ir atvertas kelias kurti visiškai naujas visuomenės, kurios potencialiai galėjo tapti turtingesnės ir taikesnės. Tačiau išmokti kurti klestinčias industrines visuomenes toli gražu nebuvo paprasta, tam prireikė daugybės brangiai kainuojančių eksperimentų ir šimtų milijonų aukų.

Vienas iš brangiai kainavusių eksperimentų buvo naujųjų laikų imperializmas. Pramoninė revoliucija prasidėjo Didžiojoje Britanijoje XVIII amžiaus pabaigoje. XIX amžiuje pramoninės technologijos ir gamybos metodai paplito kitose Europos šalyse – nuo Belgijos iki Rusijos, taip pat Jungtinėse Amerikos Valstijose ir Japonijoje. Šiuose pramonės centruose imperialistinių pažiūrų mąstytojai, politikai ir partijos skleidė idėją, kad tik stambi imperialistinė kolonijinė valstybė gali būti gyvybinga industrinė visuomenė. Jų pagrindinis argumentas – tik kolonijinė imperija gali patenkinti precedento neturinčius poreikius, nes naujosios industrinės visuomenės daug labiau priklauso nuo užsienio rinkų ir žaliavų, palyginti su savarankiškais agrarinėmis visuomenėmis. Imperialistai baiminosi, kad pramoninėms, bet kolonijų neužkariavusioms šalims agresyvesni konkurentai užkirs kelią į svarbiausias žaliavų rinkas. Kai kurių imperialistų įsitikinimu, įgyti kolonijų yra ne tik būtina, kad išliktų jų pačių valstybė, bet ir naudinga likusiai žmonijai. Jų nuomone, tik kolonijinės imperijos gali skleisti naujų technologijų palaimą vadinamajame neišsivysčiusiame pasaulyje.

Taigi, tokios pramoninės šalys kaip Didžioji Britanija ir Rusija, jau turėjusios kolonijų, jas smarkiai išplėtė, o tokios šalys kaip Jungtinės Amerikos Valstijos, Japonija, Italija ir Belgija ėmėsi jas plėsti. Aprūpintos masinės gamybos šautuvais ir artilerija, varomos garo jėga ir valdomos telegrafu, pramoninių šalių kariuomenės pasklido po pasaulį nuo Naujosios Zelandijos iki Korėjos, nuo Somalio iki Turkmėnistan. Milijonai vietinių gyventojų matė, kaip jų tradicinis gyvenimo būdas byra į šipulius po ratais karinių pajėgų, atstovaujančių kolonijinių imperijų pramonei. Prireikė daugiau nei šimtmečio kančių, kol dauguma žmonių suprato, kad kolonijinės imperijos idėja yra siaubinga ir kad yra geresnių būdų kurti industrinę visuomenę ir užsitikrinti reikiamas žaliavas bei rinkas.

Stalinizmas ir nacizmas taip pat buvo itin brangiai kainavę eksperimentai, siekę sukurti industrinę visuomenę. Tokie lyderiai kaip Stalinas ir Hitleris teigė, kad pramoninė revoliucija išlaisvino milžiniškas galias, bet suvaldyti ir visapusiškai išnaudoti jas gali tik totalitarizmas. Minėdami Pirmąjį pasaulinį karą – pirmą „totalinį karą“ istorijoje – jie įrodinėjo, kad norint išlikti pramoniniame pasaulyje reikia totaliai kontroliuoti visus politikos, visuomenės ir ekonomikos aspektus. Jie taip pat teigė, kad pramoninė revoliucija yra tarsi krosnis, išlydanti visas ankstesnes socialines struktūras su jų žmogiškaisiais trūkumais ir silpnybėmis ir suteikianti galimybę kurti tobulas visuomenes, kuriose gyvens neprilygstami antžmogiai.

Kurdami tobulą industrinę visuomenę, stalinistai ir naciai išmoko pramoniniu mastu žudyti milijonus žmonių. Traukiniai, spygliuotos vielos ir telegrafu siunčiami įsakymai buvo sujungti į neregėtą žudymo mašiną. Šiandien žvelgdami atgal dauguma žmonių bailsi tuo, ką įvykdė stalinistai ir naciai, tačiau anuomet jų drąsios vizijos hipnotizavo milijonus žmonių. 1940-aisiais daugeliui buvo lengva patikėti, kad Stalinas ir Hitleris taiko puikų pramoninių technologijų valdymo modelį, o neryžtingos liberaliosios demokratijos keliauja į istorijos šiukšlyną.

Jau vien dėl to, kad egzistavo konkuruojantys industrinės visuomenės kūrimo receptai, kilo brangiai kainavusių susidūrimų. Abu pasauliniai karai ir Šaltasis karas buvo tarsi debatai dėl tinkamo industrinės visuomenės kūrimo būdo, per kuriuos visos šalys mokėsi viena iš kitos, kartu

išbandydamos naujus pramoninio masto kariavimo būdus. Vykstant šiems debatams žuvo dešimtys milijonų žmonių, žmonija pavojingai priartėjo prie susinaikinimo.

Pramoninė revoliucija pažeidė ekologinę pusiausvyrą visame pasaulyje ir taip sukėlė rūšių nykimo bangą, ką jau kalbėti apie kitas katastrofas. Manoma, kad XXI amžiaus pradžioje kasmet išnyksta iki penkiasdešimt aštuonių tūkstančių rūšių, o bendra stuburinių gyvūnų populiacija 1970–2014 metais sumažėjo 60 procentų. Grėsmė kyla ir žmonių civilizacijos išlikimui. Kadangi vis dar nesugebame sukurti industrinės visuomenės, kuri kartu būtų ir ekologiškai tvari, dabartinės žmonių kartos išaukštinta gerovė baisesiai atsiliepia kitoms gyvoms būtybėms ir būsimoms žmonių kartoms. Galbūt ilgainiui sugebėsime – gal pasitelkę dirbtinį intelektą – sukurti ekologiškai tvarią industrinę visuomenę, tačiau kol kas Blake'o šėtonišku malūnų prisiekusių žiuri vis dar svarsto.

Jeigu kurį laiką nekreipsime dėmesio į nuolatinę žalą ekosistemai, galėsime guostis mintimi, kad galiausiai žmonės išmoko kurti geranoriškesnes industrines visuomenes. Siaubingi eksperimentai – kolonijinių imperijų užkariavimas, pasauliniai karai, genocidas ir totalitariniai režimai – davė žmonėms pamoką, kaip to *nedaryti*. Kai kas netgi teigia, kad XX amžiaus pabaigoje žmonija daugiau ar mažiau viską suprato teisingai.

Tačiau net ir tokiu atveju žinia XXI amžiui yra niūri. Jeigu žmonijai prireikė tiek daug siaubingų pamokų, kad išmoktų valdyti garo energetiką ir telegrafą, kiek jų reikės norint išmokti valdyti bioinžinerijos ir DI technologijas? Protingų mašinų, galinčių priimti sprendimus ir kurti naujas idėjas, atsiradimas reiškia, kad pirmą kartą istorijoje galia iš žmonių pereina kažkam kitam. Arbaletai, muškietos ir atominės bombos pakeitė žmogaus raumenis žudant, bet negalėjo pakeisti žmogaus smegenų sprendžiant, ką žudyti. „Little Boy“ („Mažylis“) – bomba, numesta ant Hirošimos, – sprogo su 12 500 tonų trotilo jėga, bet, kalbant apie smegenų galią, „Mažylis“ buvo bejėgis. Jis nieko negalėjo nuspręsti.

Su kompiuteriais yra kitaip. Intelektu požiūriu jie gerokai pranoksta ne tik atomines bombas, bet ir visas ankstesnes informacines technologijas, tokias kaip molio lentelės, spausdinimo presai ir radijo imtuvai. Molinėse lentelėse buvo saugoma informacija apie mokesčius, tačiau ne jos sprendė,

kokio dydžio mokesčių rinkti, o ir visai naujo mokesčio negalėjo išrasti. Spausdinimo presai kopijavo informaciją, pavyzdžiui, Bibliją, tačiau jie negalėjo nuspręsti, kokius tekstus įtraukti į Bibliją, ir negalėjo parašyti naujų šventraščio komentarų. Radijo imtuvai skleidė informaciją, pavyzdžiui, politines kalbas ir simfonijas, tačiau jie negalėjo nuspręsti, kokias kalbas ar simfonijas transliuoti, ir negalėjo jų kurti. Visus šiuos dalykus gali kompiuteriai ir netgi daugiau – jie gali sukurti naujus masinio naikinimo ginklus. Spaustuvės ir radijo imtuvai buvo pasyvūs įrankiai žmogaus rankose, o kompiuteriai jau tampa aktyviais veikėjais, kurie išvengia mūsų kontrolės ir supratimo, ir jie gali imtis iniciatyvos formuojant visuomenę, kultūrą ir istoriją.

Galbūt galiausiai rasime būdą, kaip kontroliuoti DI ir panaudoti juos žmonijos labui. Bet ar mums reikėtų pereiti dar vieną pasaulinių imperijų, totalitarinių režimų ir pasaulinių karų ciklą, kad išsiaiškintume, kaip geranoriškai panaudoti dirbtinį intelektą? XXI amžiaus technologijos yra kur kas galingesnės – ir potencialiai kur kas destruktivesnės, – palyginti su buvusiomis XX amžiuje. Vadinasi, mums reikia mažiau klysti. Galima sakyti, kad per pramoninių technologijų naudojimo pamoką XX amžiuje žmonija gavo trejetą su minusu. Vos vos išlaikė. XXI amžiuje kartelė yra daug aukščiau. Šį kartą privalome pasiekti geresnių rezultatų.